

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2002, 20:49:16 ; Search time 25.54 Seconds
(without alignments)
440.277 Million cell updates/sec

Title: US-09-664-326-23

Perfect score: 368

Sequence: 1 LTYDCTESGNLCICESGN.....PRQSHNDGDFEIPETYLQ 65

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	247	67.1	84	5	007557
2	230	62.5	84	5	007558
3	138	37.5	25	5	091WX5
4	136	37.0	25	5	091WX4
5	74.5	20.2	1331	13	090W93
6	74	20.1	248	5	096962
7	74	20.1	2352	5	061240
8	71	19.3	608	5	09V080
9	70	19.0	323	5	093NV8
10	69	18.8	315	13	09W600
11	68.5	18.6	2233	5	094711
12	68	18.5	184	2	09RB30
13	68	18.5	278	2	09XCP8
14	68	18.5	364	2	084968
15	66.5	18.1	516	3	0960Y3
16	66.5	18.1	645	5	09VTR2

17	66.5	18.1	762	13	042373	042373 brachydanio
18	66	17.9	761	4	060477	060477 homo sapien
19	66	17.9	1193	13	090819	090819 gallus gall
20	66	17.9	1242	13	090157	090157 brachydanio
21	66	17.9	2150	5	044131	044131 caenorhabdi
22	65.5	17.8	213	5	09GTV8	09GTV8 plasmodium
23	65.5	17.8	248	5	096961	096961 geodia cydo
24	65.5	17.8	594	5	024970	024970 giardia lam
25	65.5	17.8	915	4	09BXN9	09BXN9 homo sapien
26	65.5	17.8	2601	4	09BZ84	09BZ84 homo sapien
27	65.5	17.8	2601	4	096186	096186 homo sapien
28	65.5	17.8	2624	4	09BZ85	09BZ85 homo sapien
29	65.5	17.8	2624	4	096187	096187 homo sapien
30	65.5	17.8	2689	4	09BZ87	09BZ87 homo sapien
31	65.5	17.8	2689	4	096185	096185 homo sapien
32	65.5	17.8	2721	4	09BZ83	09BZ83 homo sapien
33	65.5	17.8	2721	4	096190	096190 homo sapien
34	65.5	17.8	2724	4	09BZ88	09BZ88 homo sapien
35	65.5	17.8	2724	4	096189	096189 homo sapien
36	65.5	17.8	2812	4	09BZ86	09BZ86 homo sapien
37	65.5	17.8	2812	4	096188	096188 homo sapien
38	65	17.7	558	5	017797	017797 caenorhabdi
39	65	17.7	760	11	092578	092578 rattus norv
40	65	17.7	760	11	0920P3	0920P3 mus musculu
41	65	17.7	761	11	090XL0	090XL0 mus musculu
42	65	17.7	1043	5	017644	017644 caenorhabdi
43	65	17.7	1101	5	0964D2	0964D2 entamoeba h
44	65	17.7	1661	5	077244	077244 chlorofydra
45	64.5	17.5	769	5	024971	024971 giardia lam

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	84 AA.
007557	007557			
AC	007557			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	HIRUDIN HM2 PRECURSOR (BUPREDIN).			
GN	HM2.			
OS	Hirudinaria manillensis (Buffalo leech).			
OC	Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea;			
OC	Arycnobdellida; Hirudiniiformes; Hirudinalidae; Hirudinaria.			
OX	NCBI_TaxID=6419;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 21-84.			
RX	MEDLINE=93285156; PubMed=7685281;			
RA	Scacchi E., Nitti G., Valsasina B., Orsini G., Visco C., Ferrara M.,			
RA	Saverio R.T., Sammler P.;			
RT	"Novel hirudin variants from the leech Hirudinaria manillensis. Amino			
RT	acid sequence, cDNA cloning and genomic organization.";			
RL	Eur. J. Biochem. 214:295-304(1993).			
CC	-1- FUNCTION: HIRUDIN IS A POTENT THROMBIN-SPECIFIC PROTEASE			
CC	INHIBITOR. IT FORMS A STABLE NON-COVALENT COMPLEX WITH ALPHA-			
CC	THROMBIN, THEREBY ABOLISHING ITS ABILITY TO CLEAVE FIBRINOGEN.			
CC	-1- SIMILARITY: BELONGS TO THE HIRUDIN FAMILY.			
CC	EMBL: X72786; CA51293.1; -			
DR	HSSP; P01050; 5HR.			
DR	InterPro: IPR000429; Hirudin.			
DR	Pfam: PF00713; Hirudin.1.			
DR	PRINTS; PR00777; HIRUDIN.			
DR	ProDom; PD04216; Hirudin.1.			
KW	Serine protease inhibitor; Multigene family; Signal.			
FT	SIGNAL	1	20	
FT	CHAIN	21	84	HIRUDIN HM2.
FT	DISULFID	26	34	BY SIMILARITY.
FT	DISULFID	36	48	BY SIMILARITY.
FT	DISULFID	42	57	BY SIMILARITY.
SO	SEQUENCE	84 AA;	9004 MW;	CCBA5D85E71BAF07 CRC64;

Query Match	Similarity	67.1%	Score 247	DB 5	Length 84
Best Local	Similarity	71.0%	Pred. No. 2.9e-23		
Matches	44	Conservative	6	Mismatches	10
				Indels	2
				Gaps	1
QY	1	LVYPTCTESGONLCICGESSNVGCGNKCTIGSDGSKNQCVCETGPKPDSHNDGDFEELP	60		
Db	21	VSYDTCTESGQNYCICVGSNSVGCBEKKNQGLSSG--NCCVHGEGTPKPSQTEGDFEELP	78		
QY	61	EE 62			
Db	79	DE 80			
RESULT	2				
007558					
ID	007558	PRELIMINARY	PRT	84 AA	
AC	007558				
DT	01-NOV-1998	(TREMBLrel. 08, Created)			
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)			
DE	HIRUDIN HMI	PRECURSOR.			
CN	HMI.				
OS	Hirudinaria manillensis (Bufileo leech).				
OC	Eukaryota; Metazoa; Annelida; Clitellata; Hirudina; Hirudinea;				
CC	Arynchobdellida; Hirudiniiformes; Hirudiniidae; Hirudinaria.				
OX	NCBI_Taxid=6419;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 21-84.				
RX	MEDLINE=93285156; PubMed=7685281;				
RA	Scachei E., Nitti G., Valsasina B., Orsini G., Visco C., Ferreira M.,				
RA	Sawyer R.T., Samliantos P.;				
RT	"Novel hirudin variants from the leech Hirudinaria manillensis. Amino				
RL	acid sequence. cDNA cloning and genomic organization.";				
CC	Eur. J. Biochem. 214:295-304(1993).				
CC	-I- FUNCTION: HIRUDIN IS A POTENT THROMBIN-SPECIFIC PROTEASE				
CC	INHIBITOR. IT FORMS A STABLE NON-COVALENT COMPLEX WITH ALPHA-				
CC	THROMBIN, THEREBY ABOLISHING ITS ABILITY TO CLEAVE FIBRINOGEN.				
CC	-I- SIMILARITY: BELONGS TO THE HIRUDIN FAMILY.				
DR	EMBL: X72785; CAA51292.1; -				
DR	HSSP; P01050; SHR.				
DR	InterPro; IPR000429; Hirudin.				
DR	Pfam; PF00713; Hirudin; 1.				
DR	PRINTS; PF00777; HIRUDIN.				
DR	Prodom; PD004216; Hirudin; 1.				
DR	KW Serine protease inhibitor; Multigene family; Signal.				
FT	SIGNAL	1	20		
FT	CHAIN	21	84	HIRUDIN HMI.	
FT	DISULFID	26	34	BY SIMILARITY.	
FT	DISULFID	36	48	BY SIMILARITY.	
FT	DISULFID	42	57	BY SIMILARITY.	
SO	SEQUENCE	84 AA;	8934 MW;	3AD37E7D72D4CID CRC64;	
Query Match	62.5%;	Score 230;	DB 5;	Length 84;	
Best Local Similarity	64.5%;	Pred. No. 3.7e-21;			
Matches	40;	Conservative	8;	Mismatches	12;
				Indels	2;
				Gaps	1;
QY	1	LVYPTCTESGONLCICGESSNVGCGNKCTIGSDGSKNQCVCETGPKPDSHNDGDFEELP	60		
Db	21	VSYDTCTESGQNYCICVGSNSVGCBEKKNQGLSSG--NCCVHGEGTPKPSQTEGDFEELP	78		
QY	61	EE 62			
Db	79	DE 80			
RESULT	3				
09TWS					
ID	09TWS	PRELIMINARY	PRT	25 AA	
AC	09TWS				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)			

DE	01-DEC-2001 (Tremblurel. 19, last annotation update)
DE	HIRUDIN VARIANT HV2-THROMBIN INHIBITOR.
OS	Hirudinaria manillensis (Buffalo Leech).
OC	Euryxota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
CC	Arycnobdellida; Hirudiniiformes; Hirudindae; Hirudinaria.
OX	NCBI_Taxid=6419;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=92126909; PubMed=1773002;
RA	Electricwala A., Sawyer R.T., Jones C.P., Atkinson T.;
RT	"Isolation of thrombin inhibitor from the leech Hirudinaria
RT	manillensis".;
RL	Blood Coagul. Fibrinolysis 2:83-89(1991).
DR	HSSP; P01050; IHC.
DR	InterPro; IPRO00429; Hlrudin.
DR	PRINTS; PR00777; HIRUDIN.
DR	Prodrom; PD004216; Hlrudin; 1.
SO	SEQUENCE 25 AA; 2596 MW; 5296A69273A9457D CRC64;
OY	1 LTYDTESGONLCCEGSNWCGG 25
Db	1 TTYDTESGONLCCEGSNWCGK 25
RESULT 4	
O9TWX4	PRELIMINARY; PRT; 25 AA.
AC O9TWX4;	
DT 01-MAY-2000 (Tremblurel. 13, Created)	
DT 01-MAY-2000 (Tremblurel. 13, Last sequence update)	
DT 01-JUN-2001 (Tremblurel. 17, Last annotation update)	
DE HIRUDIN VARIANT HV1-THROMBIN INHIBITOR.	
OS Hirudinaria manillensis (Buffalo Leech).	
OC Euryxota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;	
CC Arycnobdellida; Hirudiniiformes; Hirudindae; Hirudinaria.	
OX NCBI_Taxid=6419;	
RN [1]	
RP SEQUENCE.	
RX MEDLINE=92126909; PubMed=1773002;	
RA Electricwala A., Sawyer R.T., Jones C.P., Atkinson T.;	
RT "Isolation of thrombin inhibitor from the leech Hirudinaria	
RT manillensis".;	
RL Blood Coagul. Fibrinolysis 2:83-89(1991).	
DR HSSP; P01050; IHC.	
DR InterPro; IPRO00429; Hlrudin.	
DR PRINTS; PR00777; HIRUDIN.	
DR Prodrom; PD004216; Hlrudin; 1.	
SO SEQUENCE 25 AA; 2580 MW; BC50A9273A9457E CRC64;	
Query Match 37.0%; Score 136; DB 5; Length 25;	
Best Local Similarity 92.0%; Pred. No. 4,4e-10;	
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
OY 1 LTYDTESGONLCCEGSNWCGG 25	
Db 1 VVYDTESGONLCCEGSNWCGG 25	
RESULT 5	
O9OWM3	PRELIMINARY; PRT; 1331 AA.
ID O9OWM3;	
AC O9OWM3;	
DT 01-DEC-2001 (Tremblurel. 19, Created)	
DT 01-DEC-2001 (Tremblurel. 19, Last sequence update)	
DT 01-DEC-2001 (Tremblurel. 19, Last annotation update)	
DE XANTHINE DEHYDROGENASE.	
OS Poecilia reticulata (Guppy).	

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
 OC Cyprinodontiformes; Poeciliidae; Poecilia.
 OX NCBI_TaxID=8081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CAUDAL FIN;
 RA Ben J., Lim T.-M., Chan W.-K., Phang V.P.E.;
 RT "Molecular cloning of xanthine dehydrogenase gene from guppy (Poecilia
 RT reticulata).";
 RU Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY034103; AAK5969.1;
 SQ SEQUENCE 1331 AA; 146336 MW; EBBE00AA6D36D126 CRC64;

Query Match 20.2%; Score 74.5; DB 13; Length 1331;
 Best Local Similarity 32.2%; Pred. No. 1.4;
 Matches 19; Conservative 7; Mismatches 22; Indels 11; Gaps 3;

QY 2 TYTDCRES-GQNLCLCEG-----SNVCGGKNCILGSDGKNCQCVTGEGTRKPSH 51
 DB 138 TMDITQALGNCRCRCYRPIYDGCRTFCQEGNCC--QANGADCLNGSGNTNSESF 194

RESULT 6
 O96962 PRELIMINARY; PRT; 248 AA.
 AC O96962;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TETRASPANIN-CD63 RECEPTOR.
 GN CD63R.
 OS Suberites dumuncula.
 OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
 OC Hadromerida; Suberitidae; Suberites.
 OX NCBI_TaxID=55567;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mueller W.E.G., Schatton W., Wimmer W., Bohm M., Batel R., Filic Z.;
 RT "Initiation of an aquaculture of sponges for their sustainable
 RT production of bioactive metabolites in open systems.";
 RL Mol. Mar. Biol. Biotechnol. 1:569-579(1999).
 DR EMBL: Y18100; CAA77026.1;
 DR InterPro: IPR00301; Transmem_4.
 DR Pfam: PF00335; transmembrane4; 1.
 DR PRINTS: PR00259; TMFOUR.
 KW Receptor.
 SQ SEQUENCE 248 AA; 26701 MW; DB0F8D349A2E06EE CRC64;

Query Match 20.1%; Score 74; DB 5; Length 248;
 Best Local Similarity 31.3%; Pred. No. 0.26;
 Matches 21; Conservative 10; Mismatches 18; Indels 18; Gaps 4;

QY 3 YTPCTESGONL---CLCEGSNVCQGNKNCILGSDGKNCQCVTGEGTRKPSHNDG---DF 56
 DB 155 YTDYTDDEEELPPSCIDDND--GDGKNCILPISMGVNSTV-----YTDCRDSF 202

QY 57 EEIPEEY 63
 DB 203 VDFPREY 209

RESULT 7
 O61240 PRELIMINARY; PRT; 2352 AA.
 AC O61240;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HNRNOTCH PROTEIN.

GN HNRNOTCH.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 OX NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hori S., Saitoh T., Matsumoto M., Makabe K.W., Nishida H.;
 RT "Notch homologue from Halocynthia roretzi is preferentially expressed
 RT in the central nervous system during ascidian embryogenesis.";
 RU Dev. Genes Evol. 207:371-380(1997).
 DR EMBL: AB001327; BAA25571.1;
 DR HSSP: P00740; 1EDM.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_II.
 DR InterPro: IPR000800; Notch.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00008; EGF; 32.
 DR Pfam: PF00066; notch; 3.
 DR PRINTS: PR00010; EGFBL00D.
 DR PRINTS: PR01452; NOTCH.
 DR SMART: SM00179; EGF_CA; 17.
 DR SMART: SM00001; EGF_like; 9.
 DR SMART: SM00004; NL; 2.
 DR PROSITE: PS50088; ANK_REPEAT; 5.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 18.
 DR PROSITE: PS00023; EGF_1; UNKNOWN_28.
 DR PROSITE: PS01186; EGF_2; 22.
 DR PROSITE: PS01187; EGF_CA; 18.
 DR ANK repeat; Calcium-binding; EGF-like domain; Glycoprotein;
 KW Hydroxylation; Repeat.
 SQ SEQUENCE 2352 AA; 252623 MW; 13DB1C056BBD08D CRC64;

Query Match 20.1%; Score 74; DB 5; Length 2352;
 Best Local Similarity 44.4%; Pred. No. 3;
 Matches 20; Conservative 4; Mismatches 17; Indels 4; Gaps 3;

QY 2 TYTCTESGONLCLCEGSNVCQGNKNCILGSDGKNCQCVTGEGTRKPSHNDG---DF 45
 DB 977 TGTTC-ESLQN--LCTGPNCKNGSCVQTSNTVSCNCLGEGT 1018

RESULT 8
 O9YU80 PRELIMINARY; PRT; 608 AA.
 AC O9YU80;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CG10089 PROTEIN.
 GN CG10089.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abri1 J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Bendale J., Bayraktaroglu U., Beasley E.M.,
 RA Beeson K.Y., Benson P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Butris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dushin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ikegawa C.,
 RA Jajall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jammal B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003537; AAF49810.1; -.
 DR HSSP: Q16828; IMKP.
 DR FlyBase: FBgn0036369; CG10089.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR SMART: SM00195; DSPC; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SO SEQUENCE 608 AA; 66963 MW; 71EA135E5C7660D1 CRC64;

Query Match 19.3%; Score 71; DB 5; Length 608;
 Best Local Similarity 31.8%; Pred. No. 1.6;
 Matches 14; Conservative 10; Mismatches 18; Indels 2; Gaps 1;

OY 11 ONMLCEGNSVCGGKNCILGSDGKNCVTEGTPKPSHNDG 54
 DB 199 QNRDICEGN-CSRGEKCTGADVDGYNIGQGEEDGEGEG 240

RESULT 9
 O93NV8 PRELIMINARY; PRT; 323 AA.

AC O93NV8; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MAJOR SURFACE PROTEIN (FRAGMENT).
 OS Ehrlichia phagocytophila.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Anaplasma.
 OX NCBI_TaxID=948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WT 1;
 RA MEDLINE=21320744; PubMed=11427556;
 RA Lodes M.J., Mohamath R., Reynolds L.D., McNeill P., Kolbert C.P.,
 RA Bruinsma E.S., Benson D.R., Hofmeister E., Reed S.G., Houghton R.L.,
 RA Persing D.H.;
 RT "Serodiagnosis of human granulocytic ehrlichiosis by using novel

RT combinations of immunoreactive recombinant proteins.";
 RL J. Clin. Microbiol. 39:2466-2476(2001).
 DR EMBL: AF356509; AAK69699.1; -.
 FT NON-TER 1
 SO SEQUENCE 323 AA; 34533 MW; 51E0646ED4F8FDED CRC64;

Query Match 19.0%; Score 70; DB 2; Length 323;
 Best Local Similarity 29.6%; Pred. No. 1.1;
 Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;

OY 6 CRESQNLCLGSGNVCGGN-----KCLISDGKN-QCYTGEGTPKPSHNDGFEI 56
 DB 185 CSDNSSTTTSGTNSVSEGFQFIRATLKEDGSKNMTSSGTGTPTNDNAKAVA 244

OY 57 -----EIEPPE 62
 DB 245 KDLVQELTPEE 255

RESULT 10
 O9W600 PRELIMINARY; PRT; 315 AA.

AC O9W600; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE FLIK PROTEIN.
 GN FLIK.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Patel K.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96427388; PubMed=8812133;
 RA Patel K., Connolly D., Amthor H., Nose K.;
 RT "Cloning and early dorsal-axial expression of Flk1, a chick
 RT dorsoislatin-related gene: Evidence for involvement in
 RL dev. Biol. 178:327-342(1996)."
 DR EMBL: AJ238977; CAB42968.1; -.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR003645; FOIN.
 DR InterPro: IPR002350; kazal.
 DR Pfam: PF00050; kazal; 1.
 DR SMART: SM00274; FOIN; 1.
 DR SMART: SM00280; KAZAL; 1.
 SO SEQUENCE 315 AA; 35815 MW; C01889E005658A67 CRC64;

Query Match 18.8%; Score 69; DB 13; Length 315;
 Best Local Similarity 33.3%; Pred. No. 1.4;
 Matches 21; Conservative 6; Mismatches 26; Indels 10; Gaps 3;

OY 7 TSSGNLCLC-----EGSNVCGGKNCILG-SDGKNCVTEGTPKPSHNDGFEI 59
 DB 45 TEKGEPTCLICDQCKPHGRPVGSGNGKTYLNCCELHRACLIGS--KIQVDYDGCKEK 101

OY 60 PEE 62
 DB 102 KSE 104

RESULT 11
 ID O94711 PRELIMINARY; PRT; 2233 AA.
 AC O94711;

DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 51C SURFACE PROTEIN.
 OS Parametrium tetraurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;
 OC Parametrium.
 NCBI_TaxID=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92106337; PubMed=1762150;
 RA Nielsen E., You Y., Forney J.;
 RT "Cysteine residue periodicity is a conserved structural feature of
 RT variable surface proteins from Parametrium tetraurelia.";
 RL J. Mol. Biol. 222:835-841(1991).
 DR EMBL: M65164; AAA61740.1; -;
 DR Interpro: IPR002895; Parametrium_SA.
 DR Pfam: PF01508; Parametrium_SA; 25.
 SO SEQUENCE 2233 AA; 237078 MW; C064FE0AF7B873B CRC64;

Query Match 18.6%; Score 68.5; DB 5; Length 2233;
 Best Local Similarity 39.3%; Pred. No. 14;
 Matches 17; Conservative 4; Mismatches 13; Indels 9; Gaps 3;

OY 6 CTESGONICLCEGSNVCQGN---KCIIGSDGEKNQCVTGECT 45
 Q9RB30
 ID 09RB30 PRELIMINARY; PRT; 184 AA.
 AC 09RB30:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 44 KDA MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN P44-15.
 OS Ehrlichia phagocytophila.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Anaplasma.
 OX NCBI_TaxID=948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HZ;
 RX MEDLINE=99292751; PubMed=10364227;
 RA Zhi N., Ohashi N., Rikihisa Y.;
 RT "Multiple p44 genes encoding major outer membrane proteins are
 RT expressed in the human granulocytic ehrlichiosis agent.";
 RL J. Biol. Chem. 274:17828-17836(1999).
 DR EMBL: AF135261; AAD41485.1; -;
 FT NON_TER 1
 FT NON_TER 184
 SO SEQUENCE 184 AA; 19245 MW; 2BC9AD4775FD55A1 CRC64;

RESULT 12
 Q9RB30
 ID 09RB30 PRELIMINARY; PRT; 184 AA.
 AC 09RB30:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 44 KDA MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN P44-15.
 OS Ehrlichia phagocytophila.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Anaplasma.
 OX NCBI_TaxID=948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HZ;
 RX MEDLINE=99292751; PubMed=10364227;
 RA Zhi N., Ohashi N., Rikihisa Y.;
 RT "Multiple p44 genes encoding major outer membrane proteins are
 RT expressed in the human granulocytic ehrlichiosis agent.";
 RL J. Biol. Chem. 274:17828-17836(1999).
 DR EMBL: AF135261; AAD41485.1; -;
 FT NON_TER 1
 FT NON_TER 184
 SO SEQUENCE 184 AA; 19245 MW; 2BC9AD4775FD55A1 CRC64;

Query Match 18.5%; Score 68; DB 2; Length 184;
 Best Local Similarity 29.6%; Pred. No. 1;
 Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;
 OY 6 CTESGONICLCEGSNVCQGN---KCIIGSDGEKN-QCVTGECTPKPOSHNDGDF-- 56
 Db 77 CGDNGSSITITNSGANVSTGVFRDFIRATLKEGDSKNWPTSSGIGTFRKPYTNDNAKAVA 136

OY 57 -----EETPEE 62
 Db 137 KDLVQELTPEE 147

RESULT 13
 Q9XCP8 PRELIMINARY; PRT; 278 AA.
 ID Q9XCP8

AC Q9XCP8;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 44 KDA MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN P44-15.
 OS Ehrlichia phagocytophila.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Anaplasma.
 OX NCBI_TaxID=948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HZ;
 RX MEDLINE=99292751; PubMed=10364227;
 RA Zhi N., Ohashi N., Rikihisa Y.;
 RT "Multiple p44 genes encoding major outer membrane proteins are
 RT expressed in the human granulocytic ehrlichiosis agent.";
 RL J. Biol. Chem. 274:17828-17836(1999).
 DR EMBL: AF135256; AAD41480.1; -;
 DR Interpro: IPR002566; Surface_Ag_msp4.
 DR Pfam: PF01617; Surface_Ag_2; 1.
 FT NON_TER 1
 FT NON_TER 1
 SO SEQUENCE 278 AA; 29389 MW; BD5E539BB3AE16C7 CRC64;

Query Match 18.5%; Score 68; DB 2; Length 278;
 Best Local Similarity 29.6%; Pred. No. 1.6;
 Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;

OY 6 CTESGONICLCEGSNVCQGN---KCIIGSDGEKN-QCVTGECTPKPOSHNDGDF-- 56
 Db 120 CGDNGSSITITNSGANVSTGVFRDFIRATLKEGDSKNWPTSSGIGTFRKPYTNDNAKAVA 179

OY 57 -----EETPEE 62
 Db 180 KDLVQELTPEE 190

RESULT 14
 O84968
 ID 084968 PRELIMINARY; PRT; 364 AA.
 AC 084968:
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAJOR SURFACE PROTEIN-2C.
 GN MSP-2C.
 OS Ehrlichia phagocytophila.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Ehrlichieae; Anaplasma.
 OX NCBI_TaxID=948;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=USG3;
 RX MEDLINE=98339872; PubMed=9673253;
 RA Murphy C.I., Storey J.R., Recchia J., Doros-Richert L.A.,
 RA Gingrich-Baker C., Munroe K., Bakken J.S., Coughlin R.T., Beltz G.A.;
 RT "Major antigenic proteins of the agent of human granulocytic
 RT ehrlichiosis are encoded by members of a multigene family.";
 RL Infect. Immun. 66:3711-3718(1998).
 DR EMBL: AF029323; AAC31309.1; -;
 DR Interpro: IPR002566; Surface_Ag_msp4.
 DR Pfam: PF01617; Surface_Ag_2; 1.
 SO SEQUENCE 364 AA; 38806 MW; 32DA9B830A6153F2 CRC64;

Query Match 18.5%; Score 68; DB 2; Length 364;
 Best Local Similarity 29.6%; Pred. No. 2.2;
 Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;

OY 6 CTESGONICLCEGSNVCQGN---KCIIGSDGEKN-QCVTGECTPKPOSHNDGDF-- 56
 Db 226 CGDNGSSITITNSGANVSTGVFRDFIRATLKEGDSKNWPTSSGIGTFRKPYTNDNAKAVA 285

OY 57 -----EELPPE 62
 DB 286 KDLVQLRPEE 296

RESULT 15

O96UY3 PRELIMINARY; PRT; 516 AA.
 AC O96UY3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE VMA1 (FRAGMENT).
 GN VMA1.
 OS Saccharomyces sp. DH1-1A.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID-173900;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DH1-1A;
 RX MEDLINE-21486811; PubMed-11600710;
 RA Gimble F.S.;
 RT "Degeneration of a homing endonuclease and its target sequence in a
 wild yeast strain."
 RL Nucleic Acids Res. 29:4215-4223(2001).
 DR EMBL; AF389405; AAL18609.1; -.
 FT NON_TER 1
 FT NON_TER 1
 SO SEQUENCE `516 AA; 57644 MW; B41C8024561DC862 CRC64;

Query Match 18.1%; Score 66.5; DB 3; Length 516;
 Best Local Similarity 25.3%; Pred. No. 4.9;
 Matches 19; Conservative 13; Mismatches 28; Indels 15; Gaps 4;

OY 1 LYYTDTESGONICLCEGSNVCGQ---GNKCTLSGDEKNOCV---TGEGT-----P 46
 DB 24 IYVGCFAKGTNYLMADGSIETIENIEVGNK-VMGKDGPRREVYIKLPKGSETMYSYVQKS 82
 OY 47 KPOSHNDGDFFELPPE 61
 DB 83 QHRAHKSDDSSREKPE 97

Search completed: June 24, 2002, 20:51:49
 Job time: 153 sec